A potential-reduction variant of renegar's short-step path-following method for linear programming

den Hertog, Dick; Roos, C.; Terlaky, T.

Published in:
Linear Algebra and its Applications

Publication date:
1991

Link to publication

Citation for published version (APA):
A Potential-Reduction Variant of Renegar's Short-Step Path-Following Method for Linear Programming

D. den Hertog, C. Roos, and T. Terlaky *
Faculty of Technical Mathematics and Computer Science
Delft University of Technology
Mekelweg 4
2628 CD Delft, The Netherlands

Submitted by Masakazu Kojima

ABSTRACT

We propose a new polynomial potential-reduction method for linear programming, which can also be seen as a large-step path-following method. We do an (approximate) linesearch along the Newton direction with respect to Renegar's strictly convex potential function if the iterate is far away from the central trajectory. If the iterate lies close to the trajectory, we update the lower bound for the optimal value. Dependent on this updating scheme, the iteration bound can be proved to be $O(\sqrt{n} L)$ or $O(nL)$. Our method differs from the recently published potential-reduction methods in the choice of the potential function and the search direction.

1. INTRODUCTION

Since Karmarkar [6] presented his projective method for the solution of the linear programming problem in 1984, many other variants have been developed by researchers. Most of these variants can be classified into four main categories: projective methods, affine scaling methods, path-following methods, and affine potential-reduction methods.

In general, path-following algorithms start sufficiently close to the central path and follow this path closely towards the optimum, by taking very short...
steps. These characteristics of the path-following algorithms have made them unattractive in practice.

The recently developed affine potential-reduction methods don’t have these disadvantages. Many algorithms of this category can be viewed as large-step path-following algorithms. Among them are the primal-dual potential-reduction method of Ye [11] and Freund [2]. These algorithms are based on reducing the primal-dual potential function.

\[ f_Y(x, s) = q \ln x^T s - \sum_{i=1}^{n} \ln x_i s_i, \]

where \( q \) is a positive real number. It has been proved that either doing a linesearch along the primal projected steepest descent or recomputing the dual variables leads at least to a constant reduction in the potential function. The dual variables are recomputed if the primal iterate is close to the central path.

Roos and Vial [9] proposed another large-step path-following algorithm based on reductions of the primal logarithmic penalty barrier function, defined as

\[ f_{BP}(x, \mu) = \frac{e^T x - z^*}{\mu} - \sum_{i=1}^{n} \ln x_i, \]

where \( z^* \) is the optimal objective value. In this approach projected Newton steps with linesearches are taken, with the penalty parameter \( \mu \) fixed, until the iterate returns to the vicinity of the trajectory. After that, the penalty parameter is reduced by a large factor. They also showed polynomiality for this method. In essence the same approach was independently done by Gonzaga [3], in a more general way.

Gonzaga [4] also proposed another large-step path-following algorithm, which is based on the primal potential function.

\[ f_{CP}(x, z) = q \ln(e^T x - z) - \sum_{i=1}^{n} \ln x_i, \]

where \( z \) is a lower bound for the optimal value \( z^* \). When the iterate lies close to the central trajectory, the lower bound is updated by large steps, whereafter linesearches along the projected steepest-descent directions are done to return to the vicinity of the central trajectory.
In the barrier-function approach of Roos and Vial [9] and Gonzaga [3] line-searches are done along the Newton direction with respect to the logarithmic barrier function. This Newton direction coincides with the projected steepest-descent direction. In the potential-function reduction approach of Ye [11], Freund [2], and Gonzaga [4] line-searches are also done along projected steepest-descent directions. In these approaches it is impossible to prove polynomiality for the case that Newton directions are used instead of projected steepest-descent directions, because $f_1$ and $f_{c_A}$ are not necessarily convex.

In this paper we propose another large-step path-following algorithm. We deal with the linear programming problem in standard dual format. Our method is based on the following dual potential function:

$$f(y, z) = -q \ln(b^T y - z) - \sum_{i=1}^{n} \ln s_i.$$ 

In this paper we assume that $q$ is a (positive) integer, as is needed in the proof of some of our results. As one of the referees pointed out, it might be worthwhile to get rid of this assumption. Note that for $q = n$ the potential function is exactly the same as the one used by Renegar [8].

One may consider two different potential-reduction methods: a method which does line-searches along the projected steepest-descent direction and a method which uses line-searches along projected Newton directions. In [1] it is shown that the projected steepest-descent direction with respect to Renegar's potential function is

$$p' = d_{aff} + \frac{b^T y - z}{q} d_{cent},$$

where $d_{aff}$ is the dual affine scaling direction, defined as

$$d_{aff} = (AS^{-2}A^T)^{-1} b,$$

and $d_{cent}$ is the dual centering direction, defined as

$$d_{cent} = -(AS^{-2}A^T)^{-1} AS^{-1} e.$$

The same techniques as used in other potential-reduction methods can be used to develop a polynomial method based on this direction.
In this paper we will show that we can also develop a polynomial method based on the projected Newton direction with respect to Renegar's potential function. In [1] it is shown that this direction is (after reparametrizing)

\[ p = d_{\text{aff}} + \frac{(b^Ty - z)^2 / q + b^Td_{\text{cent}}}{b^Ty - z - b^Td_{\text{cent}}} d_{\text{cent}}. \]

Hence, while all other potential-reduction methods do linesearches along projected steepest-descent directions, in our method we do linesearches along projected Newton directions. We note that also in Renegar's [8] short-step path-following method a step is taken along the Newton direction \( p \), but the central path is followed very closely.

One of the referees noted that some results obtained in this paper are closely related to Nesterov and Nemirovsky's results in their monograph [17], which was unknown to the authors when we submitted this paper. The referee also observed that our results are sharper, because we deal with a specific potential function.

This paper is organized as follows. In Section 2 we describe our algorithm. Then, in Section 3, we prove some lemmas needed for the convergence analysis in Section 4. Finally, in Section 5 we show how to obtain primal feasible solutions.

**NOTATION.** Throughout the paper we use the following notation. If \( s \) denotes a vector, then the corresponding capital letter \( S \) will denote the diagonal matrix with the components of \( s \) on the diagonal. The vector \( e \) will always denote an all-one vector of appropriate length. The identity matrix will be denoted by \( I \). The vector norm \( \| \cdot \| \) will always denote the Euclidean norm \( \| y \| = (y^T y)^{1/2} \). Superscripts will be used to denote different iterates, e.g. \( y^1 \) and \( y^2 \).

2. THE ALGORITHM

We consider the dual formulation of the linear programming problem:

\[
(D) \quad \max b^Ty; \\
A^Ty + s = c, \\
s \geq 0.
\]

Here \( A \) is an \( m \times n \) matrix, and \( b \) and \( c \) are \( m \)- and \( n \)-dimensional vectors respectively. The \( n \)-dimensional vector \( y \) is the variable in which the
maximization is done. Without loss of generality we assume that all the coefficients are integer. We shall denote the length of the input data of (D) by \( L \).

We make the standard assumption that the set of optimal solutions for (D) is bounded, and that the feasible region has a nonempty interior. Additional assumptions will be made in the sequel of this paper.

It is easy to verify that \( f(y, z) \) is strictly convex on the relative interior of the feasible region. It also takes infinite values on the boundary of the feasible set. Hence it achieves a minimum value at a unique point, denoted as \( y(z) \). The necessary and sufficient Karush-Kuhn-Tucker conditions for this unique minimizing point are

\[
\begin{align*}
ATy + s &= c, \quad s \geq 0, \\
Ax &= b, \quad x \geq 0, \\
Xs &= \frac{b^Ty - z}{q},
\end{align*}
\]

where \( x \) is a \( n \)-dimensional vector. Hence \( y(z) \) lies on the central trajectory of problem (D).

For the potential function \( f(y, z) \) we can easily compute the gradient and Hessian matrix:

\[
g(y, z) := \nabla f(y, z) = -\frac{q}{b^Ty - z}b + \sum_{i=1}^{n} \frac{a_i}{c_i - a_i^Ty},
\]

and

\[
H(y, z) := \nabla^2 f(y, z) = \frac{q}{(b^Ty - z)^2}bb^T + \sum_{i=1}^{n} \frac{a_ia_i^T}{(c_i - a_i^Ty)^2}.
\]

If no confusion is possible we will write, for shortness' sake, \( g \) and \( H \) instead of \( g(y, z) \) and \( H(y, z) \).

Note that, roughly speaking, the original linear programming problem has now become a series of unconstrained optimization problems, namely \( \min f(y, z) \) with increasing lower bound \( z \). One way of solving these problems is doing linesearches along projected steepest-descent directions. This direction is simply the opposite of the gradient.

Another well-known, and more promising, method of solving unconstrained optimization problems is Newton's method, which is used in our algorithm. In Newton's method the gradient of \( f(y, z) \) at the minimum \( y(z) \)
is expanded in a Taylor series about the current iterate $y$, so that
\[ g(y(z), z) = g(y, z) + H(y, z)(y(z) - y) + \cdots. \]

Neglecting third-order and higher-order terms, let $y'$ denote the minimum of the quadratic approximation of $f(y, z)$. Then
\[ g(y, z) + H(y, z)(y' - y) = 0. \]
Hence
\[ y' - y = -H(y, z)^{-1}g(y, z). \]

We will apply linesearch along the Newton direction
\[ p(y, z) := -H(y, z)^{-1}g(y, z) = -H^{-1}g. \]

This will be repeated until the iterate is close to the central trajectory. We will use the $H$-norm $\|\cdot\|_H$ to measure closeness of points, and especially closeness to the central trajectory. The definition of this norm is as follows:
\[ \|x\|_H = \sqrt{x^T H(y, z) x}. \]

Because $H$ is positive definite, $\|\cdot\|_H$ defines a norm. We will stop linesearching along Newton directions if the following proximity criterion is satisfied:
\[ \|p\|_H \leq \epsilon, \]
where $\epsilon$ is a certain tolerance, and $p = p(y, z)$. Note that $\|p\|_H = 0$ if and only if $y = y(z)$. The same proximity criterion is used by Jarre [5]. If the proximity criterion holds, we update the lower bound as follows:
\[ z' := z + \theta(b^Ty - z), \]
where $0 < \theta < 1$. The whole process is repeated until some stopping criterion is satisfied. We note that $z'$ is really a lower bound for $z^*$, because
\[ z < b^Ty < z^*. \]
We can now describe the algorithm.

**LONG-STEP PATH-FOLLOWING ALGORITHM.**

**Input:**
- \( \theta \) is the reduction factor, \( 0 < \theta < 1 \);
- \( t \) is an accuracy parameter, \( t \in \mathbb{N} \);
- \( \epsilon \) is the proximity tolerance (we shall take \( \epsilon = \frac{1}{2} \));
- \( y^0 \) is a given interior feasible point, and \( z^0 \) is a lower bound for the optimal value, such that \( \| p(y^0, z^0) \|_{H(y^0, z^0)} \leq \epsilon \), \( z^0 \leq b^T y^0 \), and \( z^* - z^0 \leq 2L \).

\[
\begin{align*}
\text{begin} \\
y := y^0; & \quad z := z^0; \\
\text{while } & \quad b^Ty - z > 2^{-t} \text{ do} \\
\text{begin (outer step)} & \\
& \quad \text{while } \quad \| p \|_H > \epsilon \text{ do} \\
& \quad \text{begin (inner step)} \\
& \quad \quad \alpha := \arg \min_{\alpha > 0} \{ f(y + \alpha p, z) : s - \alpha A^T p > 0 \} \\
& \quad \quad y := y + \alpha p \\
& \quad \quad \text{end (inner step)} \\
& \quad z := z + \theta (b^T y - z); \\
& \quad \text{end (outer step)} \\
\text{end.}
\end{align*}
\]

In the input of the algorithm we assume that the initial point is close to the central path. It is well known in the literature that such a point can be obtained by transforming the problem; see e.g. Renegar [8]. Later on this "centering assumption" will be alleviated.

3. PRELIMINARY LEMMAS

In Section 4 we will prove that the Long-Step Path-Following Algorithm is polynomial. The next lemmas are needed to prove an upper bound for the total number of inner iterations. The lemmas are built up as follows:

- Lemma 1 states that if we do a linesearch along the Newton direction, then a sufficient decrease in the potential value can be guaranteed;
- Lemma 2 states that the sequence of iterates, obtained by doing unit steps in the Newton direction, converges quadratically to the exact center if the initial iterate fulfill the proximity criterion;
- Lemma 3 gives an upper bound for the difference in potential value of the approximately centered iterate and the exact center;
Lemma 4 states that if the lower bound is updated then the potential value increases by a constant; Lemma 5 will be used in Lemma 6, to give a relation between $b^Ty - z$ and $b^Ty(z) - z$, for the case that $y$ fulfills the proximity criterion; Lemmas 7 and 8 give some properties for $b^Ty(z) - z$.

**Lemma 1.** The decrease $Δf$ in the potential function after a linesearch along the Newton direction $p$ satisfies

$$Δf ≥ \|p\|_H - \ln(1 + \|p\|_H).$$

**Proof.** We expand $f(y + αp, z)$ in a Taylor series about $y$ as follows:

$$f(y + αp, z) = f(y, z) + αg^Tp + \frac{α^2}{2}p^THp + \sum_{j=3}^{∞} (-1)^j α^j \left( \frac{q(b^Tp)^j}{(b^Ty - z)^j} + \sum_{i=1}^{n} \frac{(a_i^Tp)^j}{(a_i^Ty - c_i)^j} \right). \quad (2)$$

We have

$$p^THp = \|p\|^2_H,$$

and, using that $g = -Hp$,

$$g^Tp = -p^THp = -\|p\|^2_H.$$

For $j = 3,4,...$ we may write

$$\left| \frac{q(b^Tp)^j}{(b^Ty - z)^j} + \sum_{i=1}^{n} \frac{(a_i^Tp)^j}{(a_i^Ty - c_i)^j} \right| \leq \left( \frac{q(b^Tp)^2}{(b^Ty - z)^2} \right)^{j/2} + \sum_{i=1}^{n} \left( \frac{(a_i^Tp)^2}{(a_i^Ty - c_i)^2} \right)^{j/2}$$

$$\leq \left( \frac{q(b^Tp)^2}{(b^Ty - z)^2} + \sum_{i=1}^{n} \frac{(a_i^Tp)^2}{(a_i^Ty - c_i)^2} \right)^{j/2}$$

$$= (p^THp)^{j/2}$$

$$= \|p\|^j_H. \quad (3)$$
Hence it follows that

\[ \sum_{j=3}^{\infty} \frac{(-1)^j \alpha^j}{j} \left( \frac{q(b^T p)^j}{(b^T y - z)^j} + \sum_{i=1}^{n} \frac{(a_i^T p)^j}{(a_i^T y - c_i)^j} \right) \]

\[ \leq \sum_{j=3}^{\infty} \frac{\alpha^j}{j} \|p\|_H^j \]

\[ = -\ln(1 - \alpha\|p\|_H) - \alpha\|p\|_H - \frac{\alpha^2}{2}\|p\|_H^2. \]

The last equality holds only if

\[ \alpha\|p\|_H < 1. \]  \tag{4} \]

Substituting all these expressions into (2), we obtain

\[ f(y + \alpha p, z) - f(y, z) \leq -\alpha(\|p\|_H^2 + \|p\|_H) - \ln(1 - \alpha\|p\|_H). \]

Hence, since \( \Delta f = f(y, z) - f(y + \alpha p, z) \),

\[ \Delta f \geq \alpha(\|p\|_H^2 + \|p\|_H) + \ln(1 - \alpha\|p\|_H). \]  \tag{5} \]

The value \( \alpha = 1/(1 + \|p\|_H) \) maximizes the right-hand side of the inequality (5). This can easily be verified by setting the derivative equal to zero. This value for \( \alpha \) also satisfies the condition (4). Replacing \( \alpha \) by this value yields the lemma.

**Lemma 2.** Let \( p^* \) and \( H^* \) be the Newton direction and the Hessian matrix at \( y^* = y + p \). If \( \|p\|_H < 1 \), then \( y^* \) is feasible and \( \|p^*\|_H < \|p\|_H^2 \).

**Proof.** Let the matrix \( \tilde{A} \) be given by

\[ \tilde{A} := (a_1 \cdots a_n a_{n+1} \cdots a_{n+q}), \]

where \( a_j := -b \) for \( n + 1 \leq j \leq n + q \). The components of the diagonal matrix \( \tilde{S} \) are defined as \( s_i \) for \( 1 \leq i \leq n \) and \( b^T y - z \) for \( n + 1 \leq i \leq n + q \).
We also introduce the matrix \( B := \tilde{A}\tilde{S}^{-1} \). Now it can easily be verified that \( H = BB^T \) and \( g = Be \). Hence, the Newton step \( p \) is determined by

\[
BB^Tp = -Be. \tag{6}
\]

We define the vector \( v \) as \( v := B^Tp \). We then have

\[
\|v\|^2 = p^TBB^Tp = \|p\|_H^2.
\]

Hence, because \( \|p\|_H \leq 1 \) we have

\[
-e \leq v \leq e. \tag{7}
\]

For the slack vector in \( y^* \) we obtain

\[
\bar{s}^* = \bar{s} - \tilde{A}^Tp = \tilde{S}(e - v) \geq 0, \tag{8}
\]

where the last inequality follows from (7). This means that \( y^* \) is feasible. From the definition of \( v \) we derive that \( v \) is in the column space of \( B^T \), which is equal to the row space of \( B \). Moreover, from (6) we derive that \( Be = -Be \). This means that \( v \) is the least 2-norm solution of the equation \( B(\bar{v} + e) = 0 \). As a consequence, \( v^* \) will be the least 2-norm solution of the equation \( B^*(\bar{v}^* + e) = 0 \). This equation is equivalent to \( \tilde{A}(\bar{s}^*)^{-1}(\bar{v}^* + e) = 0 \). Because of \( \bar{s}^* = \bar{s} - \bar{S}V \), from (8), it follows that \( v^* \) is the least 2-norm solution of

\[
B(I-V)^{-1}(\bar{v}^* + e) = 0. \tag{9}
\]

If we set \( (I-V)^{-1}(\bar{v}^* + e) \) equal to \( v + e \), then \( v^* \) satisfies (9). After some algebraic manipulations this reduces to setting \( v^* = -Vv \). Hence we have

\[
\|v^*\|^2 \leq \|\bar{v}^*\|^2 = \|Vv\|^2 \leq \|v\|^4,
\]

which proves the lemma.
Proof. Taking \( \alpha = 1 \) in (2), we get

\[
f(y + p, z) = f(y, z) + g^T p + \frac{1}{2} p^T H p
\]

\[
+ \sum_{j=3}^{\infty} \frac{(-1)^j}{j} \left( \frac{q(b^T p)^j}{(b^T y - z)^j} + \sum_{i=1}^{n} \frac{(a_i^T p)^j}{(a_i^T y - c_i)^j} \right). \tag{11}
\]

Using the same arguments as in the proof of Lemma 1, one can show that

\[
f(y, z) - f(y + p, z) \leq -\|p\|_H - \ln(1 - \|p\|_H)
\]

\[
\leq -\epsilon - \ln(1 - \epsilon)
\]

\[
\leq \frac{\epsilon^2}{2(1 - \epsilon)}, \tag{12}
\]

where the last inequality is due to Karmarkar [6]. Because \( \|p\|_H < 1 \), it follows from Lemma 2 that \( y \) lies in the region of quadratic convergence. This also means that the sequence of iterates obtained by repeatedly taking a unit step in the Newton direction converges to the exact center \( y(z) \). The lemma now follows by repeatedly applying Lemma 2 and (12):

\[
f(y) - f(y(z)) \leq \sum_{i=0}^{\infty} \frac{\epsilon^{2i+1}}{2(1 - \epsilon^{2i})}
\]

\[
\leq \sum_{i=0}^{\infty} \frac{\epsilon^{2i+1}}{2(1 - \epsilon)}
\]

\[
\leq \frac{\epsilon^2}{2(1 - \epsilon)} \sum_{i=0}^{\infty} \epsilon^{2i}
\]

\[
= \frac{\epsilon^2}{2(1 - \epsilon)(1 - \epsilon^2)}.
\]
LEMMA 4. Let $z'$ be the new lower bound, i.e. $z' = z + \theta(b^Ty - z)$, where $0 < \theta < 1$. Then

$$f(y, z') - f(y, z) = -q \ln(1 - \theta).$$

Proof. The proof is simple and straightforward. We have

$$b^Ty - z' = b^Ty - z - \theta(b^Ty - z) = (1 - \theta)(b^Ty - z).$$

Hence

$$f(y, z') - f(y, z) = q \ln \frac{b^Ty - z'}{b^Ty - z} = -q \ln(1 - \theta).$$

LEMMA 5. Let $y$ be such that $\|p\|_H \leq \epsilon < \frac{1}{2}$. Then

$$\|y - y(z)\|_{H(y(z), z)} \leq \frac{\epsilon(1 - \epsilon)(1 + 2\epsilon)}{(1 + \epsilon)(1 - 2\epsilon)}.$$

Proof. Let $y^k$ be the sequence of iterates obtained by repeatedly taking a unit Newton step $p^k$, starting from $y$, i.e. $y^{k+1} = y^k + p^k$, where $y^0 := y$, and let $H^k := H(y^k, z)$. Using the notation introduced in the proof of Lemma 2, we derive, for $i = 1, 2, \ldots, k - 1$ and any $x \in \mathbb{R}^m$,

$$\|x\|_{H^k}^2 = x^T \tilde{A} \left( \frac{\tilde{S}_k}{\tilde{S}_{k-1}} \cdots \frac{\tilde{S}_{i+1}}{\tilde{S}_i} \right)^{-2} \tilde{A}^r x$$

$$= x^T \tilde{A} \left( \frac{\tilde{S}_k}{\tilde{S}_{k-1}} \cdots \frac{\tilde{S}_{i+1}}{\tilde{S}_i} \right)^{-2} \tilde{S}_i^{-1} \tilde{A}^r x$$

$$= \left\| (I - V_{k-1})^{-1} \cdots (I - V_i)^{-1} \tilde{S}_i^{-1} \tilde{A}^r x \right\|^2$$

$$\leq \left( \frac{1}{1 - \epsilon^{2^i}} \cdots \frac{1}{1 - \epsilon^{2^{i+1}}} \cdots \frac{1}{1 - \epsilon^{2^k+1}} \right) \|x\|_{H^i}^2,$$

where the last inequality follows because $\|v^i\| = \|p^i\|_H \leq \epsilon^{2^i}$, according to
Lemma 2. In particular we have

$$\|p^i\|_{H(y(z), z)} \leq \frac{1}{\tau(e^{2i})}\|p^i\|_{H^i} \leq \frac{e^{2i}}{\tau(e^{2i})},$$

where $\tau(e) = \prod_{i=0}^{\infty} (1 - e^{2i})$. Consequently, we obtain that

$$\|y - y(z)\|_{H(y(z), z)} = \|(y^1 - y^2) + (y^2 - y^3) + \cdots \|_{H(y(z), z)}$$

(13)

$$\leq \sum_{i=0}^{\infty} \|p^i\|_{H(y(z), z)}$$

$$\leq \sum_{i=0}^{\infty} \frac{e^{2i}}{\tau(e^{2i})}$$

(14)

$$\leq \frac{\epsilon + (1 - \epsilon)e^{2\sum_{i=0}^{\infty} e^{2i}}}{\tau(\epsilon)}$$

(15)

Now we derive a simple lower bound for $\tau(\epsilon)$. It can easily be verified that $\tau(\epsilon)$ is of the following form:

$$\tau(\epsilon) = \sum_{k=0}^{\infty} \alpha_k \epsilon^k,$$

where $\alpha_k$ is either 1 or $-1$. Now, let $I_1$ be the set $I_1 = \{i: \alpha_i = 1\}$ and $I_2$ the set $I_2 = \{i: \alpha_i = -1\}$. Then

$$\tau(\epsilon) = \sum_{k \in I_1} \epsilon^k - \sum_{k \in I_2} \epsilon^k = \sum_{k=0}^{\infty} \epsilon^k - 2 \sum_{k \in I_2} \epsilon^k \geq \frac{1}{1 - \epsilon} - \frac{2\epsilon}{1 - \epsilon} = \frac{1 - 2\epsilon}{1 - \epsilon}.$$ 

The lemma follows by substituting this into (15).
Lemma 6. If \( \|y - y(z)\|_{H(y(z), z)} \leq \beta \) then

\[
\left(1 - \frac{\beta}{\sqrt{q}}\right)(b^Ty(z) - z) \leq b^Ty - z \leq \left(1 + \frac{\beta}{\sqrt{q}}\right)(b^Ty(z) - z).
\]

Proof. By definition we have

\[
\beta^2 \geq \|y - y(z)\|_{H(y(z), z)}^2
\]

\[
= (y - y(z))^T \left( q \frac{bb^T}{[b^Ty(z) - z]^2} + \sum_{i=1}^{n} \frac{a_ia_i^T}{[c_i - a_i^Ty(z)]^2} \right) [y - y(z)]
\]

\[
\geq (y - y(z))^T q \frac{bb^T}{[b^Ty(z) - z]^2} [y - y(z)]
\]

\[
= q \frac{[b^Ty - b^Ty(z)]^2}{[b^Ty(z) - z]^2}.
\]

Consequently

\[
- \frac{\beta}{\sqrt{q}} [b^Ty(z) - z] \leq b^Ty - b^Ty(z) \leq \frac{\beta}{\sqrt{q}} [b^Ty(z) - z].
\]

This implies that

\[
\left(1 - \frac{\beta}{\sqrt{q}}\right)[b^Ty(z) - z] \leq b^Ty - z \leq \left(1 + \frac{\beta}{\sqrt{q}}\right)[b^Ty(z) - z].
\]

Lemma 7. One has

\[
z^* - z \leq \left(1 + \frac{n}{q}\right)[b^Ty(z) - z].
\]

Proof. The exact center \( y(z) \) minimizes the potential function for \( z \). The necessary and sufficient conditions for this are (1). From these condi-
tions we derive that \( x(z) \) is primal feasible. Moreover, using \( z^* \leq c^T x(z) \) it follows that

\[
z^* - b^T y(z) \leq c^T x(z) - b^T y(z) = x(z)^T s(z) = \frac{n}{q} \left[ h^T y(z) - z \right].
\]

Consequently,

\[
(z^* - z) - \left[ b^T y(z) - z \right] \leq \frac{n}{q} \left[ b^T y(z) - z \right].
\]

This implies

\[
z^* - z \leq \left( 1 + \frac{n}{q} \right) \left[ b^T y(z) - z \right].
\]

**Lemma 8.** The gap \( b^T y(z) - z \) decreases monotonically if \( z < z^* \) increases.

**Proof.** The system of equations (1) determines \( y(z) \) uniquely. Differentiating this system of equations with respect to \( z \), we obtain

\[
A^T y' + s' = 0,
\]

\[
Ax' = 0,
\]

\[
Sx' + Xs' = \frac{b^T y' - 1}{q} e,
\]

where \( x', y', \) and \( s' \) denote the derivatives of \( x(z), y(z), \) and \( s(z) \) with respect to \( z \). The third equation of (16) is multiplied by \( AS^{-1} \):

\[
AS^{-1} Xs' = \frac{b^T y' - 1}{q} AS^{-1} e.
\]
This can be rewritten as

$$AX^2s' = \frac{b^Ty' - 1}{q}b.$$ 

Substituting $s' = -A^Ty'$, we get

$$y' = -\frac{b^Ty' - 1}{q} (AX^2A^T)^{-1} b.$$ 

Taking the dot product of both sides with $b$ results in

$$b^Ty' = -\frac{b^Ty' - 1}{q} b^T(AX^2A^T)^{-1} b.$$ 

Because $b^T(AX^2A^T)^{-1} b$ is positive, we conclude that $0 < b^Ty' < 1$. Consequently the derivative of $b^T(y(z) - z)$, which is equal to $b^Ty' - 1$, is negative. This proves the lemma.

4. CONVERGENCE ANALYSIS

Based on the lemmas in the previous section, we will give upper bounds for the total number of outer iterations and inner iterations. In the sequel of this section we shall assume that $q > 1$. Moreover, we shall take the proximity tolerance $\epsilon$ equal to $\frac{1}{5}$. From the previous section we derive:

(Lemma 1:) If the proximity criterion doesn't hold, then we have

$$\Delta f \geq \frac{1}{25}.$$  \hfill (17)

(Lemmas 3 and 5:) If the proximity criterion holds, then we have

$$f(y, z) - f(y(z), z) \leq \frac{1}{10},$$ \hfill (18)

and

$$\|y - y(z)\|_{l^2(y(z), z)} \leq 1,$$ \hfill (19)
and then (by Lemma 6 with $\beta = 1$),

$$
\left(1 - \frac{1}{\sqrt{q}}\right)[b^T y(z) - z] \leq b^T y - z \leq \left(1 + \frac{1}{\sqrt{q}}\right)[b^T y(z) - z]. \quad (20)
$$

**Theorem 1.** The algorithm requires

$$
K = \frac{1}{\theta} \frac{q + n}{q - \sqrt{q}} O(L)
$$

outer iterations when used to find an exact solution of the problem.

**Proof.** Let $z^k$ be the lower bound in the $k$th outer iteration, and $y^k$ the iterate at the end of $k$ outer iterations. We have

$$
\frac{z^* - z^k}{z^* - z^{k-1}} = \frac{z^* - \left[z^{k-1} + \theta(b^T y^{k-1} - z^{k-1})\right]}{z^* - z^{k-1}} = 1 - \theta \frac{b^T y^{k-1} - z^{k-1}}{z^* - z^{k-1}} \leq 1 - \theta \frac{q - \sqrt{q} \qquad q + n}
$$

The last inequality follows using Lemma 7 and then (20):

$$
z^* - z^{k-1} \leq \left(1 + \frac{n}{q}\right) \left[b^T y(z^{k-1}) - z^{k-1}\right]
$$

$$
\leq \frac{1 + n}{q} \frac{1}{1 - \frac{1}{\sqrt{q}}} (b^T y^{k-1} - z^{k-1})
$$

$$
= \frac{q + n}{q - \sqrt{q}} (b^T y^{k-1} - z^{k-1})
$$
Hence, after \( k \) outer iterations we get

\[
z^* - b^T y^k \leq z^* - z^{k+1}
\]

\[
\leq \left(1 - \theta \frac{q - \sqrt{q}}{q + n}\right)^k (z^* - z^k)
\]

\[
\leq \left(1 - \theta \frac{q - \sqrt{q}}{q + n}\right)^k (z^* - z^{0}).
\]

This means that \( z^* - b^T y^K \leq 2^{-L} \) certainly holds if

\[
\left(1 - \theta \frac{q - \sqrt{q}}{q + n}\right)^k (z^* - z^{0}) \leq 2^{-L}.
\]

Taking logarithms, this inequality reduces to

\[
-K \ln\left(1 - \theta \frac{q - \sqrt{q}}{q + n}\right) \geq L + \ln(z^* - z^{0}).
\]

Since \(-\ln(1 - v) > v\), this will certainly hold if

\[
K > \frac{1}{\theta} \cdot \frac{q + n}{q - \sqrt{q}} \left[ L + \ln(z^* - z^{0}) \right].
\]

Now using the assumption \( z^* - z^0 \leq 2^L \), made in the input of the Long-Step Path-Following Algorithm, the theorem follows.

From Lemma 6 it follows that for \( t = O(L) \), the algorithm ends up with a solution \( y^K \) such that \( z^* - b^T y^K \leq 2^{-L} \), which will give rise to an exact solution.

Now we give an upper bound for the total number of inner iterations during an arbitrary outer iteration. The approach is similar to Gonzaga's approach in [3].
Theorem 2. The total number $P$ of inner iterations during an arbitrary outer iteration satisfies

$$P \leq 3 + \frac{44q\theta}{\sqrt{q + 1}} + \frac{22q\theta^2}{1 - \theta^2}.$$

Proof. We denote the lower bound used in an arbitrary outer iteration by $z'$, while the lower bound in the previous outer iteration is denoted by $z$. The iterates during this outer iteration are denoted by $y^0, y^1, \ldots, y^p$, where $y^0$ is the iterate at the beginning of the outer iteration. Because of (17) we have

$$f(y^p, z') \leq f(y^0, z') - \frac{1}{22} P.$$  \hfill (21)

Because the lower bound was updated at the beginning of the outer iteration, we have because of Lemma 4

$$f(y^0, z') - f(y^0, z) = -q \ln(1 - \theta).$$

We also have

$$f(y^p, z') - f(y^p, z) = -q \ln \frac{b^Ty^p - z'}{b^Ty^p - z}$$

$$= -q \ln \frac{b^Ty^p - z - \theta(b^Ty^0 - z)}{b^Ty^p - z}$$

$$= -q \ln \left(1 - \theta \frac{b^Ty^0 - z}{b^Ty^p - z}\right).$$

These results are substituted into (21) to obtain

$$\frac{1}{22} P \leq f(y^0, z) - f(y^p, z) - q \ln(1 - \theta) + q \ln \left(1 - \theta \frac{b^Ty^0 - z}{b^Ty^p - z}\right).$$ \hfill (22)
Because $y^0$ is almost centered, we have because of (18)

$$f(y^0, z) - f(y(z), z) < \frac{1}{10}.$$ 

Hence

$$f(y^0, z) - f(y^p, z) < \frac{1}{10} + f(y(z), z) - f(y^p, z) \leq \frac{1}{10}.$$ 

This is substituted into (22):

$$\frac{1}{22} P \leq \frac{1}{10} - q \ln(1 - \theta) + q \ln \left(1 - \frac{b^T y^0 - z}{b^T y^p - z}\right).$$  \hspace{1cm} (23)

We have because of (20)

$$b^T y^0 - z \geq \left(1 - \frac{1}{\sqrt{q}}\right) [b^T y(z) - z].$$

We also have

$$b^T y^p - z = (b^T y^p - z') + (z' - z)$$

$$\leq \left(1 + \frac{1}{\sqrt{q}}\right) [b^T y(z') - z'] + \theta (b^T y^0 - z)$$

$$\leq \left(1 + \frac{1}{\sqrt{q}}\right) [b^T y(z') - z'] + \theta \left(1 + \frac{1}{\sqrt{q}}\right) (b^T y(z) - z)$$

$$\leq (1 + \theta) \left(1 + \frac{1}{\sqrt{q}}\right) [b^T y(z) - z],$$

where the last inequality holds because it follows from Lemma 8 that $b^T y(z') - z' \leq b^T y(z) - z$. Hence we obtain

$$\frac{b^T y^0 - z}{b^T y^p - z} \geq \frac{1 - \frac{1}{\sqrt{q}}}{(1 + \theta) \left(1 + \frac{1}{\sqrt{q}}\right)} = \frac{\sqrt{q} - 1}{(1 + \theta)(\sqrt{q} + 1)}.$$
This is substituted into (23):

\[
\frac{1}{22} P \leq \frac{1}{10} - q \ln(1 - \theta) + q \ln \left(1 - \frac{\theta(\sqrt{q} - 1)}{(1 + \theta)(\sqrt{q} + 1)}\right)
\]

\[
= \frac{1}{10} + q \ln \left(1 + \frac{\theta}{1 - \theta} - \frac{\theta(\sqrt{q} - 1)}{(1 - \theta^2)(\sqrt{q} + 1)}\right)
\]

\[
= \frac{1}{10} + q \ln \left(1 + \frac{2\theta}{\sqrt{q} + 1}\right)
\]

\[
= \frac{1}{10} + q \ln \left(1 + \frac{2\theta}{\sqrt{q} + 1}\right) + q \ln \left(1 + \frac{\theta^2}{1 - \theta^2}\right)
\]

\[
< \frac{1}{10} + \frac{2q\theta}{\sqrt{q} + 1} + \frac{q\theta^2}{1 - \theta^2}.
\]

From Theorem 1 we know that the total number of outer iterations is at most

\[
\frac{1}{\theta} \cdot \frac{q + n}{q - \sqrt{q}} O(L).
\]

Hence the total number of inner iterations during the whole process is given by

\[
\frac{q + n}{q - \sqrt{q}} \left(\frac{3}{\theta} + \frac{44q}{\sqrt{q} + 1} + \frac{22q\theta}{1 - \theta^2}\right) O(L).
\]

(24)

This makes clear that if we take \( q = O(n) \), then

- if we take \( \theta = O(1/\sqrt{n}) \), then the algorithm has an \( O(\sqrt{n}L) \) iteration bound;
- if we take \( \theta = O(1) \), then the algorithm has an \( O(nL) \) iteration bound.

The first case corresponds to a small reduction factor \( \theta \). In this case we can return to the vicinity of the central trajectory in \( O(1) \) steps, while the lower
bound must be updated $O(\sqrt{n}L)$ times. The second case corresponds to a large reduction factor $\theta$. In this case we can return to the vicinity of the central trajectory in $O(nL)$ linesearches, while the lower bound must be updated $O(L)$ times.

In the input of the algorithm we assumed that the initial point lies close to the central path. The "centering assumption" can be alleviated to $f(y^0, z^0) - f(y(z^0), z^0) \leq O(\sqrt{n}L)$ for the first case and to $f(y^0, z^0) - f(y(z^0), z^0) \leq O(nL)$ for the second case. From Lemma 1 it easily follows that these alleviations don’t affect the iteration bounds.

For the second case $[\theta = O(1)]$ this alleviation implies that the algorithm can be started from almost any interior point. We only have to assume that the initial interior feasible point $y^0$ is such that $s_i^0 \geq 2^{-O(L)}$ for each $j = 1, \ldots, n$ and that the initial lower bound $z^0$ is such that $b^Ty^0 - z^0 \geq 2^{-O(L)}[b^Ty(z^0) - z^0]$. Further, since $s_i(z^0)$ is feasible to (D), it can be written as a convex combination of basic feasible solutions. The coordinates $s_i$ of each basic feasible solution satisfy $s_i \leq 2^L$, $i = 1, \ldots, n$. Therefore

$$\sum_{i=1}^{n} \ln s_i(z^0) \leq nL. \quad (25)$$

Now due to our (very weak) assumptions and (25), we have

$$f(y^0, z^0) - f(y(z^0), z^0) = -q \ln \frac{b^Ty^0 - z^0}{b^Ty(z^0) - z^0} - \sum_{i=1}^{n} \ln s_i^0 + \sum_{i=1}^{n} \ln s_i(z^0) \leq (q + n)O(L).$$

Consequently, for $q = O(n)$ we have $f(y^0, z^0) - f(y(z^0), z^0) \leq O(nL)$, which means that the algorithm can be initiated with $y^0$.

5. OBTAINING PRIMAL FEASIBLE SOLUTIONS

The algorithm proposed and analysed in the previous sections works on the dual formulation (D). In each iteration the dual variable $y$ is feasible. In some applications it is necessary to obtain feasible solutions for the primal problem

$$(P) \quad \min c^Tx:$$

$$Ax = b,$$

$$x \geq 0.$$
In this section we will show that, if the proximity criterion holds, then primal feasible solutions can be obtained. Moreover, we will give a lower and an upper bound for the corresponding duality gap. We will use the notation introduced in the proof of Lemma 2.

**Theorem 3.** Let \( w := c + e \). If \( \|p\|_H \leq 1 \) then

\[
\begin{align*}
\sum_{i=1}^{n} \frac{w_i}{s_i} a_i &= q \frac{w_{n+1}}{b^Ty - z} b,
\end{align*}
\]

is primal feasible.

**Proof.** From (6) we derive that \( Bw = 0 \). From the definition of \( B \)

\[
\sum_{i=1}^{n} \frac{w_i}{s_i} a_i = q \frac{w_{n+1}}{b^Ty - z} b,
\]

because all \( w_{n+i} \) are equal for \( 1 \leq i \leq q \). Consequently, defining \( x_i \) as

\[
\begin{align*}
x_i := \frac{b^Ty - z}{qs_i w_{n+1}} w_i, & \quad 1 \leq i \leq n,
\end{align*}
\]

it follows that \( Ax = b \). For the feasibility of \( x \) we also have to verify \( x \geq 0 \).

It is easy to verify that

\[
w^T w = w^T e = e^T e - \|p\|_H^2 = n + q - \|p\|_H^2.
\]

Let \( \tau := \|p\|_H \); then it follows that \( w \) lies on the sphere

\[
w^T w = n + q - \tau^2
\]

and in the hyperplane

\[
w^T e = n + q - \tau^2.
\]

Now we prove that the minimal value of \( w_i \), denoted \( w_{\min} \), is greater than or equal to 0 if \( \tau \leq 1 \). Using the Kuhn-Tucker theory, it can easily be verified
that the minimal value $w_{\text{min}}$ occurs if $n + q - 1$ components of the vector $w$ are equal, and one component is equal to $w_{\text{min}}$. Using (27) and (28) it is easy to verify that $w_{\text{min}}$ satisfies

$$(n + q)w_{\text{min}}^2 - 2(n + q - \tau^2)w_{\text{min}} + (n + q - \tau^2)(1 - \tau^2) = 0.$$  

From this it is obvious that $w_{\text{min}} < 0$ and $\tau < 1$ gives a contradiction. Consequently, if $\tau < 1$, then $x$ is primal feasible.

**Theorem 4.** If $\|p\|_H \leq 1$ and $x$ is defined by (26), then the duality gap $x^Ts$ satisfies

$$\frac{n}{q} (b^Ty - z) \frac{\sigma - \tau \sqrt{\sigma/n}}{\sigma + \tau \sqrt{\sigma/q}} \leq x^Ts \leq \frac{n}{q} (b^Ty - z) \frac{\sigma + \tau \sqrt{\sigma/n}}{\sigma - \tau \sqrt{\sigma/q}},$$

where $\tau := \|p\|_H$ and $\sigma := 1 - \frac{\tau^2}{n + q}$.

**Proof.** From the definition of $x_i$ we obtain

$$x_is_i = \frac{b^Ty - z}{qw_{n+1}} w_i.$$  

Consequently, for the duality gap we derive

$$x^Ts = (b^Ty - z) \frac{\sum_{i=1}^n w_i}{qw_{n+1}}.$$  

Now we will derive an upper bound for

$$\chi(w) := \frac{\sum_{i=1}^n w_i}{qw_{n+1}}.$$  

From (27) and (28) it follows that $w$ belongs to the $n + q - 1$-dimensional
sphere with center $\sigma e$ and radius $\tau \sqrt{\sigma}$, where $\sigma = 1 - \tau^2/(n + q)$. This means that
\[
\sum_{i=1}^{n} \left( w_i - \sigma \right)^2 + q \left( w_{n+1} - \sigma \right)^2 = \sigma \tau^2. \tag{29}
\]

To obtain an upper bound for $\chi(w)$, we maximize $\chi(w)$ subject to (29). This maximum is certainly less than the maximum of $\chi(w)$ subject to
\[
\sum_{i=1}^{n} \left( w_i - \sigma \right)^2 \leq \sigma \tau^2,
\]
\[
q \left( w_{n+1} - \sigma \right)^2 \leq \sigma \tau^2. \tag{30}
\]

From (30) we easily derive that $\sum_{i=1}^{n} w_i \leq n(\sigma + \tau \sqrt{\sigma/n})$ and $w_{n+1} \geq \sigma - \tau \sqrt{\sigma/q} > 0$. Using these bounds we obtain
\[
\chi(w) \leq \frac{n(\sigma + \tau \sqrt{\sigma/n})}{q(\sigma - \tau \sqrt{\sigma/q})}.
\]

In the same way it can be verified that
\[
\chi(w) \geq \frac{n(\sigma - \tau \sqrt{\sigma/n})}{q(\sigma + \tau \sqrt{\sigma/q})}.
\]

Hence, the theorem follows.

REFERENCES


11 Y. Ye, A Class of Potential Functions for Linear Programming, Dept. of Management Sciences, Univ. of Iowa, Iowa City, 1988.

Received 10 February 1990; final manuscript accepted 12 July 1990